

# METHOD FOR SCREENING INPUT IMAGE

Publication number: JP2000207565

Publication date: 2000-07-28

Inventor: RENATO KURESHU

Applicant: HEWLETT PACKARD CO

Classification:


- International: G06T1/00; G06K9/00; G06T7/00; G06T1/00; G06K9/00; G06T7/00; (IPC1-7): G06T7/00; G06T1/00

- European: G06K9/00F1

Application number: JP20000001317 20000107

Priority number(s): US19990229052 19990111

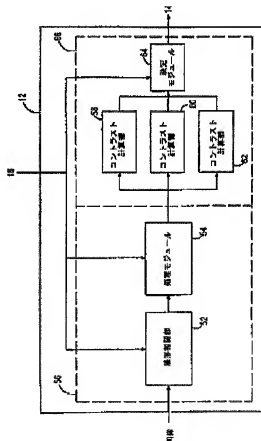
Also published as:

 US6463163 (B1)

Report a data error here

## Abstract of JP2000207565

**PROBLEM TO BE SOLVED:** To improve a detection function for a target image such as a face in a speed and accuracy by performing screening by a candidate selecting part, selecting and eliminating many input images that do not include a face in the area and subsequently processing only a candidate area by a face detecting part. **SOLUTION:** A candidate selecting part 12 includes a linear correlating part 52 and a processing module 54 and they forms a linear matched filter 56. The part 12 also includes contrast calculating parts 58, 60 and 62 and a decision module 64 and they forms a nonlinear filter 66. The filter 56 selects many interim candidate areas that seem to include a face an input image according to certain reference. The nonlinear filter examines the selected interim candidate areas and abandons undesirable candidate areas according to another reference. A face detecting part 14 processes residual areas as a candidate area. The associated operation of the filters 56 and 66 increases the selection accuracy of an optimum candidate area by the part 12.



Data supplied from the esp@cnet database - Worldwide

**Family list****2** family members for: **JP2000207565**

Derived from 2 applications

[Back to JP2000207565](#)**1 METHOD FOR SCREENING INPUT IMAGE****Inventor:** RENATO KURESHU**Applicant:** HEWLETT PACKARD CO**EC:** G06K9/00F1**IPC:** G06T1/00; G06K9/00; G06T7/00 (+5)**Publication info:** **JP2000207565 A** - 2000-07-28**2 System and method for face detection using candidate image region selection****Inventor:** KRESCH RENATO (IL)**Applicant:** HEWLETT PACKARD CO (US)**EC:** G06K9/00F1**IPC:** G06T1/00; G06K9/00; G06T7/00 (+4)**Publication info:** **US6463163 B1** - 2002-10-08

---

Data supplied from the **esp@cenet** database - Worldwide

(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-207565

(P2000-207565A)

(43) 公開日 平成12年7月28日 (2000.7.28)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	ページコード (参考)
G 0 6 T 7/00		G 0 6 F 15/70	3 3 0 Z
1/00		15/62	3 8 0

審査請求 未請求 請求項の数 1 OL (全 18 頁)

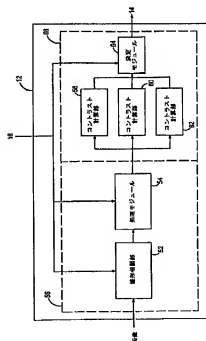
(21) 出願番号	特願2000-1317(P2000-1317)	(71) 出願人	398038580 ヒューレット・パッカード・カンパニー HEWLETT-PACKARD COMPANY アメリカ合衆国カリフォルニア州パロアルト ハノーバー・ストリート 3000
(22) 出願日	平成12年1月7日 (2000.1.7)	(72) 発明者	レナト・クレシュ イスラエル、34334、ハイファ、ディズレーリ 48エー
(31) 優先権主張番号	0 9 / 2 2 9 0 5 2	(74) 代理人	100081721 弁理士 岡田 次生
(32) 優先日	平成11年1月11日 (1999.1.11)		
(33) 優先権主張国	米国 (US)		

## (54) 【発明の名称】 入力画像を予備選択する方法

## (57) 【要約】

【課題】 速度と正確さにおいて顔などのターゲット画像の検出機能を向上させる。

【解決手段】 画像の中に目標とする顔などのターゲット画像を潜在的に含む候補領域を選択する候補選択回路(12)を用いる顔検出システムであり、候補選択回路(12)は、線形整合フィルタ(56)と非線形フィルタ(66)で構成され、これらは連続して動作して入力画像から候補画像を選び出す。



## 【特許請求の範囲】

【請求項1】入力画像が関心のあるターゲット画像パターンを含むかどうかを決定するために入力画像を予備選択する方法であって、

上記入力画像を受信するステップと、

上記入力画像と、上記関心のあるターゲット画像パターンの代表である予備選択された基準データとの相関をとり、上記入力画像と上記予備選択された基準データとの相関を示す画像データをもつ相関画像を引き出すステップと、

上記相関画像における局所極大を検出することに基づいて上記入力画像の潜在的候補領域を識別することを含み、上記関心のあるターゲット画像パターンを含む領域を持つ候補領域を選ぶステップと、を含む入力画像の予備選択方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】この発明は、一般的には、画像パターン検出および認識に関し、とりわけ、顔の検出のための装置および方法に関するものである。

## 【0002】

【従来の技術】画像を含む膨大な量のデジタル情報、現在、インターネットその他の電子的データベースを介して利用可能である。テキスト情報と異なり、これらのデータベースからターゲットオブジェクトであるピクチャを含む画像のコンテンツベースでの検索はやりがいのある作業である。記録された画像の中で、ターゲットオブジェクトであるピクチャが、通常、同じでないことがこの作業に困難を生じる。さらに、1つの画像から取り出されたターゲットオブジェクトが、他の画像から取り出された他のターゲットオブジェクトと向きおよびサイズが異なることがある。

【0003】顔検出技術は部分的には進んでおり、上記困難さを克服できる画像検索システムの開発の支援になっている。顔検出は、入力画像の中に顔のピクチャが存在するかどうかを決定するプロセスであり、在ると判定した場合には、入力画像の中で顔の位置を正確に決定する。顔検出部は、入力画像をスキャンし、その画像の中からサイズに関係なく取り出される顔を検出するように設計されている。顔検出には、「ニューラル・ネットワークに基づく」手法と、「アイゲンフェース(固有の顔; eigenface)」手法の2つの検出した取り組み方がある。

【0004】名前が示唆するように、ニューラル・ネットワークに基づく手法は、入力画像中に顔を検出するのにニューラル・ネットワークを使用する。ニューラル・ネットワークに基づく手法の基本的な考えは、1つのM×M画像ブロックを入力とし、その画像の中に顔が検出されたか否かを示す2値の出力を出力するようにニューラル・ネットワークを設計することである。ニュー

ラル・ネットワークは、トレーニング用画像ブロックから成る大きなデータベースを使ってトレーニングをする。トレーニング用画像ブロックは、顔画像と顔以外の画像の混在したものである。トレーニング用画像ブロックは、ニューラル・ネットワークに入力される前に独特の処理を施される。前処理には、画像ブロックの直流成分を除去することや、画像ブロックを正規化することが含まれる。ニューラル・ネットワークがトレーニングした後、トレーニング用画像ブロックと同様に前処理されたM×M画像ブロックの入力画像が、オンライン検出手順の中でニューラル・ネットワークによって分析可能となる。

【0005】Henry A. Rowley, Shumeet Baluja及びTakeo Kanadeによる「Neural Network-Based Face Detection」(IEEE論文集、第20巻、No.1、ページ23-28、1998年1月)にニューラル・ネットワークに基づく顔検出システムについての記述がある。Rowleyほかによる顔検出システムは、標準的なニューラル・ネットワークに基づく顔検出手法の修正版を使っている。特に、Rowleyその他による顔検出システムは、以前に顔を含むと誤判断されたトレーニング用画像ブロックを使い「ブートストラップ」法によってさらにニューラル・ネットワークをトレーニングさせるものである。上記論文は、ブートストラップ手法がニューラル・ネットワークを十分にトレーニングさせるに必要なトレーニング用画像ブロックの数を減らし、と主張している。それに加えて、Rowleyその他による顔検出システムは、処理手順の中で画像ブロックから最適線形関数を除去することによる画像ブロックにおける照度差の無効化を含んでいる。

【0006】アイゲンフェース手法は、1組のトレーニング用顔に近づくために、最適な線形基底または「アイゲンフェース」と呼ばれる主要な固有ベクトル部分を計算することを含む。そして、これらの基底ベクトルは入力画像の中に顔を検出する整合フィルターの畳み込み核として使われる。Moghaddamiほかによる米国特許No. 5,710,833は、アイゲンフェース(eigenface)手法によって画像の中に特定の特徴を検出し、認識する装置について記述している。Moghaddamiほかによる装置では、主要な固有ベクトル成分だけでなく、全ての固有ベクトルを使用する。全ての固有ベクトルを使用することは、複数の特徴を持つ複雑なオブジェクトを検出する装置の正確さを増すことを意図としている。

## 【0007】

【発明が解決しようとする課題】従来の顔検出システムはそれらの意図された目的の範囲ではよく動作するが、速度と正確さにおいて顔検出機能を向上させる顔検出装置と顔検出方法が求められている。

## 【0008】

【課題を解決するための手段】顔の検出装置および顔検出用入力画像の予備選択の方法は、潜在的に顔のピクチャ

ヤを含む候補領域を選択する候補選択部を使用する。候補選択部は、候補領域が顔を含むかどうか確認する動作に関連する顔検出部に続いて行う。予備選択は候補選択部によって行われ、その領域に顔を含まない多くの入力画像が選別除去される。その後、候補領域だけが顔検出部によって処理されるので、候補選択部による処理が、顔検出部によって行われなければならない計算量を減少させることになる。

【0009】好適な実施例において、候補選択部は線形整合フィルタと非線形フィルタを内蔵し、それらは直列に動作して入力画像から候補領域を選択する。線形整合フィルタは、顔テンプレートと比較して類似度の高い画像パターンを持つ画像領域を選択する。線形整合フィルタは線形相関器と処理モジュールを内蔵している。線形相関器はフィルタ核を使って入力画像に対して線形相関処理を行い、相関画像を取り出す。フィルタ核は顔テンプレートのひとつの数値表現である。フィルタ核はフィルタ核生成部によってトレーニング期間中或いは非顔検出期間中に計算される。線形相関処理は離散コサイン変換(DCT)ドメインで行われるのが好ましいが、他の取り組み方も可能である。相関画像は、それから処理モジュールによって調べられる。処理モジュールは、決定ルールを使って入力画像から暫定候補領域を選択するように構成される。決定ルールの出す指示は、相関画像において局所極大の近傍に位置し、かつ、画素相関値が相関閾値より大きい画像領域だけを選択せよ、というものである。暫定候補領域は、その後、非線形フィルタへ送られる。

【0010】非線形フィルタは、暫定候補領域が候補領域とみなされるべきかどうかを決める働きをする。非線形フィルタは、暫定候補領域の或る領域の範囲内でコントラスト値を調べ、そして、人間の目の特性を持つコントラストパターンを採る。この領域での高いコントラスト値は、画像領域が顔を含んでいる確度の高さと同値である。実施例において、非線形フィルタは3つのコントラスト計算部と1つの決定モジュールを内蔵する。コントラスト計算部は、画像領域の特定の上部セグメントのためのコントラスト値計算を行う。第1のコントラスト計算部は、画像領域の上半分のセグメントについてコントラスト値を計算する。第2のコントラスト計算部は、画像領域の4象限のうちの上2つについてコントラスト値を計算する。こうして、第1と第2のコントラスト計算部は暫定候補領域の上部50%のために供される。第3のコントラスト計算部は上半分のセグメントの構成部分である3つの近接するセグメント、すなわち暫定候補領域の上部33%についてのコントラスト値を計算する。これらのコントラスト値は、決定モジュールに送られる。第2と第3のコントラスト計算部で計算されたコントラスト値はコントラスト閾値と比較される。これらの値がコントラスト閾値を上回る場合、その画像領域は

候補領域とみなされて顔検出部へ送られる。代りの構成では、これら3つの計算部が1つの計算部で具体化される。

【0011】顔検出部では、候補領域の中に顔を検出する手法として、ニューラル・ネットワークに基づく手法、アイゲンフェース手法、或いは、その他の既知の手法が利用できる。好適な実施例では、顔検出部はRowleyほかによる顔検出システムであり、これはオリジナルな顔検出方式を使っている。最も好適な実施例では、顔検出部は、高速版のオリジナルな顔検出方式を使ったRowleyほかによる顔検出システムである。顔検出部は、候補選択部から候補領域を受けとり、1つ以上の候補領域が顔を含んでいるかどうかを決定する働きをする。顔検出部での判定結果は、ディスプレイ画面上で表示され、そこでは検証された候補領域がオリジナル入力原画像上にアウトラインの形を重ね合わされることによって識別される。

【0012】違ったサイズの顔を検出するために、顔検出システムは入力画像の尺度を修正する画像スケラーを内蔵している。画像スケラーは入力画像を受け取り、それを逐次縮小して候補選択部へ送る。最初に送られる入力画像は原画像のサイズのままでよい。好適な実施例においては、画像スケラーは入力画像を1.2の比率で縮小する。しかし、別の比率が用いられてもよい。

【0013】このシステムは、候補選択部の線形整合フィルタ用にフィルタ核を準備するフィルタ核生成部を有する。このフィルタ核生成部は、顔画像のサンプルを集めた大きなデータベースに基づいてフィルタ核を計算するように構成されている。このフィルタ核生成部はシステムがオンライン動作、即ち、顔検出処理工程中は動作しない。その代わりに、フィルタ核の計算は、オフライン動作、即ち、トレーニング工程中で実行される。フィルタ核は顔検出処理工程以前に計算される。

【0014】このフィルタ核生成部は、1つの平均化ユニット、1つのDCT処理部、1つのマスクア、及び、1つの逆離散コサイン変換(IDCT)処理部を内蔵する。フィルタ核の計算は平均化ユニットに顔画像のサンプルを集めた固定サイズで大きなデータベースが入力されると開始する。顔画像は、8×8画素の画像であることが望ましい。平均化ユニットは顔画像を平均化して平均化画像とし、それをDCT処理部へ送る。DCT処理部は平均化された画像を空間ドメインからDCTドメインへ変換する。変換された画像は、その後、マスクア(masker、マスク部)へ送られる。マスクアは変換された画像から、直流、照明および雑音の周波数成分を除去する。次に、その平均化された画像はIDCT処理部によって変換され空間ドメインに戻される。この結果の画像がフィルタ核である。このフィルタ核は、線形整合フィルタがこのフィルタ核があるまでシステム・メモリに記憶される。このフィルタ核は線形相関処理で使われることによって、入力画像が

ら直流、照明および雑音の影響に関連した成分も除去する。

【0015】別の実施例では、トレーニング用顔画像は最初に空間ドメインからDCTドメインへの変換、マスク、空間ドメインへの変換をされた後、フィルタ核を取り出すために平均化される。この実施例では、DCT処理部は最初にトレーニング画像を受け取る。DCT処理部はその後、受け取った各トレーニング用顔画像を空間ドメインからDCTドメインへ変換する。次に、マスクが変換された顔画像から直流、照明および雑音の成分を廃棄する。この処理された顔画像はDCT処理部によって空間ドメインへ変換される。この顔画像はその後、平均化ユニットによって平均化されフィルタ核が取り出される。

#### 【0016】

【発明の実施の形態】本発明による1つの顔検出システム10が図1に示されている。顔検出システム10は、画像スケーラ11、候補選択部12、画像メモリ13、顔検出部14、フィルタ核生成部16、メモリ18及びディスプレイ装置20を含んで構成される。顔検出システム10は、他の既知の顔検出システムと同様、採取された顔のサイズや向きの違いに関係なく入力画像をいろいろな尺度で調べることによって、入力デジタル画像の中から顔の検出を行う。しかし、従来の顔検出システムと違って、顔検出システム10は候補選択部12で実行される候補選択方式を用いている。この候補選択方式は、潜在的に顔を含む入力画像の中の1つ以上の領域を選択するために、入力画像を予備選択することを含んでいる。従って、入力画像の中の選択されなかった領域は顔を含みそうな候補ではないとして選別、除外される。選択された領域だけが更に顔検出部14で処理され、顔を含む画像であるかどうか判定される。この選別工程は候補選択部12で実行され、顔検出部14で実行しなければならない計算の量を減少させる。顔検出システム10は顔を検出するために特別に設計されたものであるが、このシステムは違った画像パターンや複数種の画像パターンの検出用に容易に修正することができる。

【0017】このシステム10の候補選択部12は、顔検出部14のほか、画像スケーラ11、画像メモリ13、フィルタ核生成部16、及びメモリ18と組んで動作する。画像スケーラ11は、画像中の異なるサイズの顔をこのシステム10が検出できるように入力画像を修正する。画像スケーラ11は入力画像を受け取り、それを逐次縮小して画像メモリ13へ送る。最初に送られる入力画像は原画像のサイズのままでよい。好適な実施例においては、画像スケーラ11は入力画像を1.2の比率で縮小する。画像メモリ13の中のスケーリングされた画像は全て、候補選択部12及び顔検出部14によってアクセス可能である。従って、候補選択部12が1つのスケーリングされた画像をアクセスしている間に顔検出部14がもう1つのスケーリ

ングされた画像をアクセスすることが可能である。

【0018】フィルタ核生成部16は、フィルタ核を生成して、それをメモリ18経由で候補選択部12に与える。メモリ18は独立した構成品として描かれているが、画像メモリ13と一体のものとしてもよい。フィルタ核は顔テンプレートと数値で表現したものであり、候補選択部12が選択動作を実行するときに用いられる。フィルタ核生成部16については、図6を参照しながら更に説明する。メモリ18には、候補選択部12及び顔検出部14で使われるデータが、フィルタ核も含めて、記憶される。顔検出部14は、候補領域が顔を含むかどうかの判定に応じて出力信号を発生する。その出力信号はディスプレイ装置20が受け取る。ディスプレイ装置20は計算機モニタであればよい。その出力信号が、顔検出部14による候補領域における顔検出が肯定的であることを示す場合は、ディスプレイ画面で、入力原画像上にその候補領域のアウトラインが重畳表示される。別の実施例では、ディスプレイ装置20に代って、例えば、顔認識モジュール、顔計数モジュール、或いは、保安システムのような処理モジュールがその出力信号を受け取り、さらに処理を行う。

【0019】顔検出部14は通常のものであり、選択された領域の中で顔を検出する手法であれば、ニューラルネットワークに基づく手法でも、アイゲンフェース(eigen face)手法その他既知のものでも使える。好適な実施例では、顔検出部14はニューラルネットワークに基づくものである。最も好適な実施例では、顔検出部14はRowleyほかによる顔検出システムであり、高速版のオリジナルな顔検出方式を使ったものである。以下、Rowleyほかによる方式の高速版について内容を説明する。Rowleyほかによる顔検出方式のオリジナル版は、高速版も同様、"Neural Network-Based Face Detection"と表題のつけられた出版物に記述されている。この出版物をここに引用する。顔検出部14には、多数ある既知の顔検出方式の中の1つを使えばよいのであるが、ここでは、最も好適な実施例、即ち、Rowleyほかによる高速版の方式を取上げて説明する。

【0020】顔検出システム10における候補選択部12と顔検出部14の全般的な動作について図2を参照しながら説明する。図2には、顔検出システム10の流れに沿った各点での入力画像22が示されている。左端の入力画像22は、顔検出システム10の候補選択部12に入力される前の画像を示している。中央の入力画像22は、候補選択部12で処理された後の画像を示している。右端の入力画像22は、顔検出部14で処理された後の画像を示している。ここでの入力画像22は、入力原画像をスケーリング処理したものであるとする。入力画像22は領域24、26、28、30、32、34、36、38、40、42、44および46に分割されている。この領域24-46は潜在的に顔が存在する領域を表わしている。顔検出システム10によってチェックされる入力画像22内の領域には領域24-46とオーバラップする

領域もある。しかし、簡単化のため、入力画像22のオーバラップした領域は無視した。

【0021】まず、入力画像22は候補選択部12によって画像メモリ13から読み出される。矢印48は候補選択部12によるこの動作が行われたことを示す。入力画像22を受け取った後、候補選択部12は画像22の各領域24-46を調べ、潜在的に顔を含む領域を選択する。候補選択部12は特定の領域を候補領域であると判定するために必要な基準を用いる。候補選択部12が用いるこの基準については後で説明する。領域32と36が顔を含まれているとき、候補選択部12が、例えば、図2の中央の画像22において太い枠で示したように、領域26、32及び36を選んだとする。

【0022】次に、画像22は顔検出部14でさらに処理される。矢印50は顔検出部14によるこの動作が行われたことを示す。顔検出部14は候補領域のみ、即ち、候補選択部12が選んだ領域26、32及び36のみを調べる。顔検出部14は既知の顔検出手法を実行して、領域26、32及び36に含まれているかどうか判定する。要するに、顔検出部14は、候補選択部12が行う顔を含む領域の選択の確かさを確認する。図の右端の画像22に示すように、顔検出部14は領域32と36が顔を含まないという肯定的な判定を下す。領域26に対して顔検出部14は、その領域が顔を含まないという否定的な判定を下す。この例においては、候補選択部12と顔検出部14の連動作によって顔を含まない2つの領域が正しく識別された。候補選択部12が無ければ、顔検出部14はどの領域が顔を含まれているかを判定するために画像22の全ての領域について調べなければならない。しかし、候補選択部12が、顔を含まない領域として殆どの領域を既に選別、除外しているので、顔検出部14が調べなければならないのは3つの領域26、32及び36だけである。このため、顔検出部14で行われる計算量は減少する。

【0023】候補選択部12の内部構成を図3に示す。候補選択部12には、1つの線形整合フィルタ52と1つの処理モジュールが含まれ、これらは線形整合フィルタ56を形成している。候補選択部12にはまた、コントラスト計算部58、60及び62と、決定モジュール64が含まれ、これらは非線形フィルタ66を形成している。線形整合フィルタ56は、入力画像の中で顔を含まないような多数の暫定候補領域を成る基準に従って選択する。それから、選ばれた暫定候補領域は非線形フィルタ66によって調べられ、望ましくない暫定候補領域が別の基準に従って廃棄される。残った領域が候補領域として顔検出部14でさらに処理される。フィルタ56と66の連動作は、候補選択部12が最適な候補領域を選択する確度を増す。

【0024】線形整合フィルタ56は、線形整合フィルタの手法を実行することにより、入力画像から暫定候補領域を選択する。線形整合フィルタの手法は、線形畳み込みを行うステップと決定するステップとを含む。線形畳

み込みを行うステップは“線形相関”とも云われ、線形相関部52により成る $N \times N$ フィルタ核を使って入力画像を対象にして実行される。対象となる入力画像は入力原画像にスケールリングの施されたものである。 $N \times N$ フィルタ核はフィルタ核生成部16で計算され、メモリ18に記憶されたものである。線形畳み込みを行うステップは、離散コサイン変換(DCT)ドメインか、空間ドメインか、或いはその他適当なドメインにおいて行うことができる。

【0025】好適な実施例においては、線形畳み込みを行うステップは、DCTドメインで実行される。DCTドメインは他のドメインに比べて線形畳み込みを実行するための計算の複雑さが少ない。例えば、線形畳み込みは、通常、空間ドメインにおけるよりもDCTドメインのような周波数ドメインにおけるほうが高速に実行できる。その上、普通の画像ブロックのDCT係数の多くは零か或いは零に近似的ものである。従って、これらのDCT係数を含む算術演算は避けることができる。更に、入力画像が、JPEG、MPEGまたはその他のDCTに基づいた圧縮規格によるものである場合、空間ドメイン用のデータを得るには逆DCT演算が必要となるのに対して、DCT係数は容易に得られる。線形相関部52で実行されるDCTドメインにおける線形畳み込みについて、これから数学的な説明を行う。線形畳み込み操作は入力画像を相関画像に変換する。最も好適な実施例では、線形畳み込みを行うステップは、計算量を減らす暗黙的な手法に従って、離散正弦変換(DST)係数を使ったDCTドメインにおいて実行される。その手法についての数学的な説明を後で行う。

【0026】変換された相関画像は処理モジュール54に送られる。処理モジュール54は、入力画像の領域の中から整合フィルタの決定ルールを満たしたものを暫定候補領域として選択する。処理モジュール54は、相関画像において局所極大を含み、かつ、固定閾値 $T_1$ を超え画面素相関値を持つ領域のみが選択されるように設定されている。局所極大というのは、相関画像上の位置であって、周りの位置の相関値よりも大きな相関値を持つところを指す。閾値 $T_1$ はメモリ18に記憶されていて、処理モジュール54からの要求があったとき、処理モジュール54に与えられる。夫々の暫定候補領域は $8 \times 8$ 画素の領域であることが望ましく、選択された暫定候補領域内の各画素は $T_1$ より大きな相関値を持つ。さらに、暫定候補領域としては、その中心に局所極大が位置している領域が選択されなければならない。選択された暫定候補領域は非線形フィルタ66へ送られる。

【0027】各暫定候補領域に対して非線形フィルタ66は、その暫定候補領域の予め定められた部分領域内で局所的なコントラスト値の分析を行う。予め定められた部分領域は暫定候補領域内で目が存在することが期待される位置に対応する。暫定候補領域のうちコントラスト閾値より大きな局所的コントラスト値を持つものは目を含

んでいと推定され、候補領域として識別される。暫定候補領域のうちコントラスト閾値と等しいか若しくは小さい局所的コントラスト値を持つものは目を含んでいないと推定され、廃棄される。

【0028】非線形フィルタ66のコントラスト計算部58、60及び62は暫定候補領域内の予め定められた部分領域内で局所的コントラスト値の計算を行う。図4にこれらの部分領域を示す。暫定候補領域68が $N \times N$ 画素の領域、ここでは $N=8$ 、として描かれている。また、図4にはいろいろなサイズの6つの部分領域 $A_0, A_1, A_2, A_3, A_4$ 及び $A_5$ が描かれている。これらの部分領域 $A_0, A_1, A_2, A_3, A_4$ 及び $A_5$ は次式で定義される。ここで、 $W(i, j) \leq i, j \leq N-1$ は暫定候補領域68を表わす。

【0029】

【数1】

$$\begin{aligned} A_0 &\triangleq \{(i, j) | 0 \leq i \leq (N/2)-1, 0 \leq j \leq (N/2)-1\} \\ A_1 &\triangleq \{(i, j) | 0 \leq i \leq (N/2)-1, 0 \leq j \leq (N/2)-1\} \\ A_2 &\triangleq \{(i, j) | 0 \leq i \leq (N/2)-1, N/2 \leq j \leq N-1\} \\ A_3 &\triangleq \{(i, j) | 0 \leq i \leq [N/4]-1, 0 \leq j \leq [3N/8]-1\} \\ A_4 &\triangleq \{(i, j) | 0 \leq i \leq [N/4]-1, [5N/8] \leq j \leq N-1\} \\ A_5 &\triangleq \{(i, j) | 0 \leq i \leq [N/4]-1, [3N/8] \leq j \leq [5N/8]-1\} \end{aligned}$$

【0030】コントラスト計算部58は部分領域 $A_0$ についてコントラスト値を計算する。部分領域 $A_0$ についてコントラスト値は $m_0-m_6$ であって、 $m_0$ は部分領域 $A_0$ 内におけるグレースケールの最大値を表わし、 $m_6$ は部分領域 $A_0$ 内におけるグレースケールの最小値を表わす。何れの $m_k$ 及び $m_6$ も次式を使って求められる。

【数2】

$$\begin{aligned} M_k &\triangleq \max_{0 \leq i \leq N-1} W(i, j) \\ m_k &\triangleq \min_{0 \leq i \leq N-1} W(i, j) \end{aligned}$$

【0031】同様に、コントラスト計算部60は部分領域

$$\min \{M_1-m_1, M_2-m_2, M_3-m_3, M_4-m_4, M_5-m_5\} > T_0,$$

$$\text{where } T_0 \triangleq \min \{(M_0-m_0)/4, T_N\}.$$

【0033】 $T_0$ は決定モジュール64に与えられるコントラスト閾値である。 $T_0$ の値はメモリ18に記憶されている。決定モジュールから要求されたときに決定モジュール64に与えられる。上記基準を満足する暫定候補領域だけが候補領域として選ばれる。選ばれた候補領域は顔検出部14へ送られる。コントラスト計算部58、60及び62は、図3では別々の構成品として示されているが、1つの構成品で実現してもよい。

【0034】準好適な実施例では、候補選択部12は線形整合フィルタ56だけを含む。この実施例においては、線形整合フィルタ56によって選ばれた暫定候補領域は、直接顔検出部14へ送られる。別の準好適な実施例では、候補選択部12は非線形フィルタ66だけを含む。この実施例においては、非線形フィルタ66が人力画像の各 $N \times N$ 領域を、それぞれスケールリング・レベルにおいて、走査することになる。そして、コントラスト基準に合致する全ての $N \times N$ 領域が候補領域として選ばれ、顔検出部14

$A_1$ 及び $A_2$ について夫々コントラスト値 $m_1-m_3$ 及び $m_2-m_3$ を計算する。コントラスト計算部62は残りの部分領域 $A_3, A_4$ 及び $A_5$ について夫々コントラスト値 $m_4-m_5, m_4-m_5$ 及び $m_5-m_5$ を計算する。

【0032】これらの部分領域について計算されたコントラスト値は、暫定候補領域68が確定候補領域として選択すべきものかどうかを決定モジュール64が判定するのに用いられる。決定モジュール64は、部分領域 $A_0, A_1, A_2, A_3, A_4$ 及び $A_5$ について計算された各コントラスト値が全て $T_0$ より大きいときに限り、その暫定候補領域が候補領域であると判定する。これを数式で表現すると、

【数3】

へ送られる。

【0035】図5には顔検出部14の構成が示されている。顔検出部14は事前処理部（preprocessor、プリプロセッサ）70とニューラル・ネットワーク72を含む。事前処理部70は候補選択部12の決定モジュール64に接続され、選ばれた候補領域を受け取る。事前処理部70は各候補領域に対する画像処理のなかで、領域が含まれている直流成分の除去、領域の正規化、及び、適切でない照明によって生じた画質の悪さの補正を行う。次に、事前処理を受けた候補領域はニューラル・ネットワーク72へ送られる。ニューラル・ネットワーク72は候補領域が顔を含んでいるかどうかの判定をする。ニューラル・ネットワーク72は全ての候補領域に関する上記判定に応じて1つの出力信号を生成する。その後、その出力信号はディスプレイ装置へ送られる。

【0036】図1に示す最も好適な実施例では、画像スケラー11、候補選択部12、顔検出部14及びフィルタ核



生成部16は1つのコンピュータ・プログラムの中に組み込まれ、マイクロプロセッサ(図示せず)で実行される。この実施例における顔検出方式はRowleyほかによるニューラル・ネットワークに基づく顔検出方式の高速版を修正したものである。従って、顔検出システム10で用いる顔検出方式を説明する前にRowleyほかによる顔検出方式について説明する。

【0037】Rowleyほかによる顔検出方式は、まず、入力画像の30 x 30画素の領域(画像ブロック)に対する候補ニューラル・ネットワーク(NN)を各スケールにおいて適用する。ここで、画素領域の中心は、iとjを10の倍数とする位置(i,j)にあるものとする。即ち、画像ブロックの中心が、10画素の間隔で離れた格子上の点から取

られている。候補NNは顔が含まれると信じる領域を選択する。次に、位置NNが用いられ、それぞれの選ばれた領域における顔のより明確な位置を推定する。位置NNは、要するに、最初に選ばれた領域から変位した新しい領域を選ぶ。最後に、確認NNが選ばれた新しい領域に適用され、新しい領域が含まれているかどうかを確認する。Rowleyほかによる顔検出方式の高速版の要点を以下に示す。記述はCに似た疑似コードでなされており、“original\_face\_detection\_routine()”及び“apply\_NN’s(i,j)”として識別される。

【0038】

【表1】

```
original_face_detection_routine() {
    calculate a pyramid of downscaled versions
        of the original image (downscale ratio = 1.2);
    for (t = 0; t < number_of_levels_in_pyramid; t++) {
        Define "downscale_image" as the level t of the pyramid;
        for (l = 0; l < height_of_downscale_image; l+=10) {
            for (j = 0; j < width_of_downscale_image; j+=10) {
                apply_NN's(t,l,j);
            }
        }
    }

    apply_NN's(t,l,j) {
        Pre-process the block centered at position (l,j)
            (remove DC and illumination components, and normalize);
        Apply the Candidate NN at position (l,j);
        If (selected) {
            Apply the Location NN to obtain the
                fine-tuned candidate position (x,y);
            Apply the Verification NN's at position (x,y);
            If (selected) {
                Declare (x,y) as the center of a face at scale t;
            }
        }
    }
}
```

【0039】顔検出システム10の顔検出方式は、Rowleyほかによる方式の部品を選択的に使うもので、以下に示すようにCに似た疑似コードで表現され、“combined\_f

ace\_detection\_routine()”で呼び出される。

【0040】

【表2】

```

combined_face_detection_routine() {
    Calculate a pyramid of downscale versions
    of the original image (downscale rate = 1.2).

    /* For all the levels of the pyramid, except for the last  $\Delta t$  ones,
    use the candidate selection scheme of the candidate selector 12 as a pre-filter for
    the NN's. */
    for (i = 0; i < number_of_levels_in_pyramid -  $\Delta t$ ; i++) {
        Calculate the DCT-domain convolution of the level i +  $\Delta t$  of the pyramid.
        Define "downscale_image" as the level i of the pyramid.
        for (j = 0; j < height_of_downscale_image; j++) {
            for (k = 0; k < width_of_downscale_image; k++) {
                If (i,j) is selected by the matched filtering decision rule)
                or If (i,j) is selected by the non-linear filter)
                apply_NNs(i,j,k); /* see pseudo-code for Rowley et al. scheme */
            }
        }
    }

    /* For the last  $\Delta t$  levels of the pyramid, use the original scheme. */
    for (i = number_of_levels_in_pyramid -  $\Delta t$ ; i < number_of_levels_in_pyramid; i++) {
        Define "downscale_image" as the level i of the pyramid.
        for (j = 0; j < height_of_downscale_image; j++) {
            for (k = 0; k < width_of_downscale_image; k++) {
                apply_NNs(i,j,k); /* see pseudo-code for Rowley et al. scheme */
            }
        }
    }
}

```

【0041】システム10の顔検出方式は、原方式が使われている最後の $\Delta t$ に対して以外全てのスケール・レベルについて用いられる。Rowleyほかによる方式とシステム10の方式の間の差違は上記擬似コードにおいて明確である。最後の $\Delta t$ を除けば、DCTドメインでの畳み込みを各スケール・レベルに対して1回行っていることも差違に含まれる。さらに、システム10ではスケールリングされた入力画像の各画素に対して内部ループが実行されるのに対して、Rowleyほかによる方式においては内部ループは画像ブロックの中心の画素に対してのみ実行される。最後に、線形整合フィルタ56の処理モジュール54によって実行されるDCTドメイン整合フィルタ決定基準、および、非線形整合フィルタ66の決定モジュール64で実行される局所コントラスト基準に基いた画素のみがニューラル・ネットワーク(NN)でチェックされる。

【0042】最も好適な実施例では、DCTドメイン線形畳み込み及び非線形フィルタのための核のサイズ $N_k$ は8である。また、 $\Delta t$ は3である。確認NNに対する核のサイズが20であり、縮小スケールリング比が1.2(従って、 $8 \times 1.2^5$ が約20になる)であることを考慮すると、理論的には、 $\Delta t$ は5であるべきである。しかし、シミュレーションで最良の結果を得たのは $\Delta t=3$ であった。さらに、線形整合フィルタ56の処理モジュール54で使われる閾値 $T_k$ は3,000。そして、非線形フィルタ66の決定モジュール5043】図6に、フィルタ核生成部16の構成を示す。

す。フィルタ核生成部16は、線形整合フィルタ56の線形相関部52で実行される線形畳み込み処理に使われるフィルタ核を計算する。フィルタ核は、フィルタ核生成部に入力される数多くのトレーニング用(トレーニング)顔画像を使って、フィルタ核生成部によって算出される。算出される核は、線形相関処理に使われる際に入力画像の各ブロックについて直流、照明及びノイズによる影響を補正できるように設計される。フィルタ核の計算は、顔検出期間中(オンライン)でなく、トレーニング期間中(オフライン)に行うのが好ましい。顔検出期間中というのは、顔検出システム10が入力画像に対して、その画像の中に描かれている可能性のある人間の顔を検出する動作を行っている期間のことである。フィルタ核の計算はフィルタ核生成部16によって、ただ1回だけ実行されればよい。

【0044】フィルタ核生成部16は、平均化ユニット74、離散コサイン変換(DCT)部76、マスク78及び逆離散コサイン変換(IDCT)部80を含んでいる。平均化ユニット74はトレーニング用顔画像を受け付ける。顔画像は $8 \times 8$ 画素の画像群である。平均化ユニット74は受け付けた画像を平均化して平均化画像を取り出すように設計されている。その後、DCT部80が平均化画像を空間ドメインからDCTドメインへ変換する。マスク78は、図7に示すようにマスク82を使って変換された画像から直流、照明及びノイズの成分を除去する動作を行う。マスク82はマスク78で除去される周波数成分を示している。除去さ

れる周波数成分は、図7において太線で囲われたブロックで示されている。ブロック(0,0)と(0,1)は、夫々、直流と照明の成分を表わしている。ブロック(4,7)、(5,6)、(5,7)、(6,5)、(6,6)、(6,7)、(7,4)、(7,5)、(7,6)及び(7,7)で形作られる下方の三角形の部分はノイズ成分を表わしている。

【0045】平均化画像のマスクが完了すると、IDCT部80がそのマスクされた画像を元の空間ドメインに逆変換する。その結果得られる画像がフィルタ核である。得られたフィルタ核はメモリ18に送られ、候補選択部12の線形相関部52からの要求があるまでそこに記憶される。フィルタ核生成部16によって実行される数値的処理については後で説明する。

【0046】1つの代替構成として、フィルタ核生成部16は次のように再編される。まず、トレーニング用顔画像が空間ドメインからDCTドメインに変換され、マスクされ、再び空間ドメインへ戻され、そしてフィルタ核を取り出すために平均化される。この代替構成では、DCT部76はトレーニング用画像を受ける立場にある。DCT部76は受けたトレーニング用画像の夫々を空間ドメインからDCTドメインへ変換する。変換された顔画像はマスク78へ送られる。マスク78は、図7に示すマスク82を使って変換された画像から直流、照明及びノイズの成分を除去し、処理後の顔画像をIDCT部80へ送る。IDCT部80はマスク処理された顔画像を元の空間ドメインへ戻す変換を行う。平均化ユニット74はIDCT部80につながっており、マスク処理された顔画像を空間ドメインで受け取る。この顔画像は平均化ユニット74によって平均化された後、メモリ18へ送られる。

【0047】別の代替構成では、フィルタ核生成部16は顔検出システム10に含まれない。この実施例では、顔検出システム10の外部にあるフィルタ核生成部16によってフィルタ核が計算される。そして、計算されたフィルタ核はメモリ18へ送られて記憶される。

【0048】顔検出のために入力画像の中で顔を潜在的に含む候補領域を選択するための、入力画像の予備選択について、図8及び図9参照しながら説明する。図8は、トレーニング期間中にフィルタ核を引き出す処理工程の流れ図である。図8は、顔検出処理中に、候補領域を選択するべく入力画像の予備選択を行う処理工程の流れ図である。まず、フィルタ核を引き出す処理工程について説明する。ステップ100では、オンライン期間中に、多くのトレーニング用顔画像がフィルタ核生成部16に入力される。次に、ステップ102では、顔画像が平均化ユニットによって平均化され、平均化画像が引き出される。平均化画像は次のステップ104で、DCT部76によって空間ドメインからDCTドメインへ変換される。ステップ106では、変換された画像から直流、照明およびノイズに対応した周波数成分がマスク78によってマスクされ除去される。つぎのステップ108では、マスク処理された画像

がDCT部によって空間ドメインへ逆変換される。その結果得られる画像がフィルタ核である。ステップ110では、フィルタ核はメモリ18に記憶され、顔検出処理動作によって、顔検出システム10の候補選択部12からの要求があるまで待つ。

【0049】代替構成では、トレーニング用顔画像はマスクが行われる直前まで平均化は実行されない。この実施例では、トレーニング用顔画像は最初にDCT部76によって空間ドメインからDCTドメインへ変換される。次に、変換された顔画像はマスク78でマスクされた後、IDCT部によって空間ドメインへ逆変換される。平均化ユニット74はマスク処理された画像を空間ドメインにおいて平均化して、フィルタ核を引き出す。最後に、フィルタ核はメモリ18に記憶される。この代替構成は効率的ではないが、得られるフィルタ核は前記実施例で得られるフィルタ核と同じである。

【0050】候補領域を選ぶための入力画像の予備選択を行うオンライン処理工程について説明する。図9において、ステップ112で、入力画像が顔検出システム10に受け取られる。次に、ステップ114で、入力画像は線形整合フィルタ56の線形相関部52によってフィルタ核との線形相関を取られ、相関画像が引き出される。この線形相関はDCTドメインにおいて取られるのが望ましい。ステップ116では、入力画像の中でその相関画像が局所極大を持つ画像領域が線形整合フィルタ56の処理モジュール54によって選択される。ステップ118では、選択された画像領域が処理モジュール54によって次のチェックを受け、選択された画像領域の各画素の相関値が或る相関閾値を超えているかどうかの判定を受ける。この閾値基準を満足する画像領域が次へ進むことを許される。

【0051】ステップ120では、画像領域の種々の部分領域についてのコントラスト値の計算を含めて、非線形フィルタ66による選択された各画像領域のグレースケール特性が分析される。この種々の部分領域は各画像領域内の上半分に限定されていることが望ましい。次に、ステップ122では、非線形フィルタ66の決定モジュール64によって特定画像領域の各コントラスト値が或るコントラスト閾値と比較される。コントラスト閾値を超えるコントラスト値を持った画像領域のみが候補領域であるとみなされる。この候補領域は、顔を含んでいるかどうかを確認するための処理を、顔検出部14など通常のどのような顔検出部によっても受けることができる。使用される顔検出部のタイプによってはステップ114-122が入力画像のいろいろな縮小版に対して繰り返される。と云うのは、候補領域が置ったサイズの顔を含んでいる可能性があるからである。

【0052】準好適な実施例では、ステップ114、116及び118は予備選択処理工程から省かれる。この実施例では、相関画像が局所極大を持つ画像領域だけが分析されるのではなく、非線形フィルタ66によって受け付けた画

像の全ての画像領域が分析され、コントラスト値が引き出される。コントラスト閾値を超えるコントラスト値を持つ画像領域が、相関値に関係なく、候補領域であるとみなされる。別の準好適な実施例では、ステップ120及び122は顔検出処理工程から省かれる。この実施例では、相関閾値を超える相関値を持つ画像領域が候補領域であるとみなされる。画像領域のコントラスト値はこの実施例には無関係である。

#### 【0053】数学的根拠

線形整合フィルタ56の線形相関部52及びフィルタ核生成部16の動作を説明するために数学的な説明を行う。はじめに、入力画像が、重なり合わない  $N \times N$  のブロックに分割される。フィルタ動作の各ステップにおいて、これら画像ブロック内の近接する4つのブロックがとりあげられ、 $x_i, i=1, \dots, 4$  で表わされる。1つの  $(N+1) \times (N+1)$  の

$$r(n, m) = \sum_{p=0}^{N-1} \sum_{q=0}^{N-1} y(p, q) \cdot x(p+n, q+m), \quad 0 \leq n, m \leq N. \quad \text{式1}$$

【0055】上記の式は下記の行列表現と等価である： 【数5】

$$r = Y_1 Y_1 + Y_2 Y_2 + Y_3 Y_3 + Y_4 Y_4, \quad \text{式2}$$

【0056】ここで、

【数6】

$$\tilde{x}_i, i = 1, \dots, 4$$

は、 $x_i$  の列スタック (column stack) ベクトル表現で

$$Y_1[n(N+1)+m, pN+q] = \begin{cases} y(q-m, p-n), & \text{if } p-N < n \leq p \text{ and } q-N < m \leq q \\ 0, & \text{otherwise.} \end{cases} \quad \text{式3}$$

行列  $Y_1, i=2, 3$  及び  $4$  も同様のフォーマットである。行列  $Y_2$  は次式で与えられる：

$$Y_2[n(N+1)+m, pN+q] = \begin{cases} y(q-m, p+N-n), & \text{if } p < n \leq p+N \text{ and } q-N < m \leq q \\ 0, & \text{otherwise.} \end{cases} \quad \text{式4}$$

【0057】行列  $Y_3$  は次式で与えられる：

$$Y_3[n(N+1)+m, pN+q] = \begin{cases} y(q+N-m, p-n), & \text{if } p-N < n \leq p \text{ and } q \leq m \leq q+N \\ 0, & \text{otherwise.} \end{cases} \quad \text{式5}$$

【0058】最後に、行列  $Y_4$  は次式で与えられる：

$$Y_4[n(N+1)+m, pN+q] = \begin{cases} y(q+N-m, p+N-n), & \text{if } p < n \leq p+N \text{ and } q \leq m \leq q+N \\ 0, & \text{otherwise.} \end{cases} \quad \text{式6}$$

【0059】目標は、DCTドメインを通して  $r$  の計算をする、効果的な他の方法を得ることである。略率的な方式

ブロック  $r$  が下記のように計算される。それから、ブロック  $r$  の左上端から  $N$  行、 $N$  列で構成される  $N \times N$  のブロックが、フィルタされた画像の1つの対応するブロックを形成する。このブロックも重なり合わない  $N \times N$  のブロックである。上記は、出力 (フィルタされた) 画像の全てのブロックが計算されるまで、入力画像の4つの近接するブロックの夫々のグループに対して繰り返される。

【0054】フィルタ核生成部16で用意された  $N \times N$  フィルタ核  $y$  が与えられると、線形整合フィルタ56の線形相関部52は  $y$  による  $x$  の線形畳み込みをとることによって得られる  $(N+1) \times (N+1)$  のブロック  $r$  を計算するべく動作する：

【数4】

あり、また、 $Y_i, i=1, \dots, 4$  は、要求された畳み込み動作を実行する  $(N+1)^2 \times N^2$  の行列である。行列  $Y_1$  は次式で与えられる：

【数7】

【数8】

【数9】

【数10】

によって得られる解は以下の式によって表わされる：  
【数11】

$$r = C_1^{-1} \{Y_a \bar{X}_a^c + Y_b \bar{X}_b^c + Y_c \bar{X}_c^c + Y_d \bar{X}_d^c\} \quad \text{式7}$$

$$X_a^c \triangleq X_1^c + \phi X_2^c + X_3^c \phi + \phi X_4^c \phi, \quad \text{式8}$$

$$X_b^c \triangleq X_1^c - \phi X_2^c + X_3^c \phi - \phi X_4^c \phi, \quad \text{式9}$$

$$X_c^c \triangleq X_1^c + \phi X_2^c - X_3^c \phi - \phi X_4^c \phi, \quad \text{式10}$$

$$X_d^c \triangleq X_1^c - \phi X_2^c - X_3^c \phi + \phi X_4^c \phi, \quad \text{式11}$$

ここで、 $C_1^{-1}$ はタイプ1-e(下記定義による)の逆DCT、 $X_i$ 、 $i=1, \dots, 4$ は夫々ブロック $x_i$ のタイプ2-e(下記定義による)のDCT係数、 $Y_i$ 、 $i \in \{a, b, c, d\}$ は $(N+1)^2 \times N$ 核行列(下記により特徴づけられた)、及び、行列 $\phi$ は対角線変調行列(diagonal modulation matrix)で、 $\text{diag}\{(-1)^p\}$ 、 $p=0, \dots, N-1$ を表す。1つの行列に対して $\phi$ の左掛け算は、その偶数行に符号反転を生じ、右からの掛け算はその偶数列に符号反転を生ずることに注意のこと。

【0060】DCT変換の定義づけ、及び、上記核行列の特徴づけを行う前に、式7～式10で要点を示され、線形相関部52で実行される暗黙的なフィルタ方式についての要約を以下に示す：

【0061】1.各ブロック $x_i$ 、 $i=1, \dots, 4$ はタイプ2-eの2

-D DCTによって変換される。

2.変換されたブロックは、その後、2-Dバタフライ(式8～式11)の手段によって結合される。このバタフライ処理中に行列 $X_i^c$ 、 $i=2, 3, 4$ は、それらの偶数の列及び/又は行で符号が反転することに要注意。

3.バタフライ処理された行列は、列スタックされ、それぞれに核行列が予め掛けられ、そして合計される。

4.最後に、タイプ1-eの逆DCTが生成されたブロックに対して適用され、そして、相関ベクトルが得られる。

【0062】ブロック $x$ に関するタイプ2-eの順DCT、 $X^c$ をここで定義する。 $0 \leq p, q \leq N-1$ に対して：

【数12】

$$X^c(p, q) \triangleq \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} Y(n, m) \cos \left( \frac{np \left( n + \frac{1}{2} \right)}{N} \right) \cos \left( \frac{nq \left( m + \frac{1}{2} \right)}{N} \right), \quad \text{式12}$$

【0063】さらに、スタック列ベクトルに適用されたときタイプ2-eの2D DCTを実行する $N \times N$  DCT行列 $C_1$ を次のように定義する：

【数13】

$$C_1^{-1} \{Y^c(n, m)\} \triangleq \sum_{p=0}^N \sum_{q=0}^N \frac{c(p)c(q)}{N\sqrt{2}} Y^c(p, q) \cos \left( \frac{pn}{N} \right) \cos \left( \frac{qm}{N} \right), \quad \text{式14}$$

ここで、 $c(0)=c(N)=1/2$ 及び $c(k)=1, \dots, N-1$ 。上記定義の如く、タイプ1-eの逆DCTは $(N+1) \times (N+1)$ のブロックの入力と出力を持つことに要注意。さらに、スタック列ベクトル $X^c$ に対して適用されたときタイプ1-eの逆2D DCTを実行する $(N+1) \times (N+1)$ 逆DCT行列 $C_1^{-1}$ は $C_1^{-1} \cdot X^c = x$

$$C_n \cdot X = X^c.$$

式13

【0064】ブロック $Y^c$ に関するタイプ1-eの逆DCTをここで定義する。 $0 \leq n, m \leq N$ に対して：

【数14】

と定義する。結局、行列 $C_1$ は $C_1^{-1}$ の逆数で与えられるタイプ1-eの順2D DCT変換を表す。

【0065】核行列 $Y_i$ 、 $i \in \{a, b, c, d\}$ は次の式で特徴づけられる。

【数15】

$$Y_2 = \frac{1}{4} C_1 \left\{ Y_1 C_{11}^{-1} + Y_2 C_{11}^{-1} (1 \otimes \phi) + Y_3 C_{11}^{-1} (\phi \otimes 1) + Y_4 C_{11}^{-1} (\phi \otimes \phi) \right\}, \quad \text{式15}$$

$$Y_3 = \frac{1}{4} C_1 \left\{ Y_1 C_{11}^{-1} + Y_2 C_{11}^{-1} (1 \otimes \phi) - Y_3 C_{11}^{-1} (\phi \otimes 1) - Y_4 C_{11}^{-1} (\phi \otimes \phi) \right\}, \quad \text{式16}$$

$$Y_4 = \frac{1}{4} C_1 \left\{ Y_1 C_{11}^{-1} - Y_2 C_{11}^{-1} (1 \otimes \phi) + Y_3 C_{11}^{-1} (\phi \otimes 1) - Y_4 C_{11}^{-1} (\phi \otimes \phi) \right\}, \quad \text{式17}$$

$$Y_5 = \frac{1}{4} C_1 \left\{ Y_1 C_{11}^{-1} - Y_2 C_{11}^{-1} (1 \otimes \phi) - Y_3 C_{11}^{-1} (\phi \otimes 1) + Y_4 C_{11}^{-1} (\phi \otimes \phi) \right\}, \quad \text{式18}$$

ここで、丸に×印はクロネッカー積を示し、1はN×N識別行列を示す。

【0066】DCT及びIDST係数(ここでは示されていない)の関数として与えられる核行列 $Y_i, i \in \{a, b, c, d\}$ を分析すると、潜在的な方式により特に対称性があり、かつ/又は、データがまばらであるときに、計算量の削減が図られるという結論が得られる。対称性というのは、最初のフィルタ核が空間ドメインにおいて水平方向、かつ/又は、垂直方向で対称であることに関係する。両方向に対称性がある場合、この方式は、どちらか1軸にある場合に比べてほぼ倍の速さになる。顔テンプレートには通常水平方向の対称性があることを記しておく。データがまばらであるということは、入力画像のDCTブロックがまばらであること、即ち、その係数の殆どがゼロである状態をいう。これは、自然な画像にとって典型的な状態であることを記しておく。さらに、少数の非ゼロ係数は、通常、DCTブロックの左上象限に存在することを記しておく。従って、この場合、核行列のほんの一部、即ち、入力ブロックの左上象限に存在するDCT係数を掛ける行列の一部、しか演算に使われていない。これが満たされれば、演算量はほぼ4分の1になる。

【0067】この発明は、例として次の実施形態を含む。

【0068】1. 入力画像(22)が関心のあるターゲット画像パターンを含むかどうかを決定するために入力画像を予備選択する方法であって、上記入力画像を受信するステップと(112)、上記入力画像と、上記関心のあるターゲット画像パターンの代表である予備選択された基準データとの相関をとって相関画像を引き出すステップと、該相関画像は、上記入力画像と上記予備選択された基準データとの相関を示す画像データを持っているステップと、上記相関画像の中に局所極大を検出することによる潜在的候補領域の識別(116)を含み、上記関心のあるターゲット画像パターンを含む確度を持つ入力画像の候補領域(26, 32及び36)を選ぶステップと(48)、を含む入力画像の予備選択方法。

【0069】2. 上記1記載の方法において、上記候補領域(26, 32及び36)を選択するステップ(48)が、更に、上記潜在的候補領域のグレースケール特性を分析する(1

20)ことによって上記潜在的候補領域を選択するステップを含み、このステップは、上記選択された候補領域がターゲット画像パターンを含む確度を高めるために、上記グレースケール特性に関して予め定められた基準を満たさない上記潜在的候補領域を廃棄することを含む方法。

【0070】3. 上記2記載の方法において、上記潜在的候補領域を選択するステップが、更に、各潜在的候補領域の部分領域について計算されたコントラスト値と、上記関心のあるターゲット画像パターンが人間の顔であるとしての或るコントラスト閾値とを比較するステップ(122)を含む方法。

【0071】4. 上記1, 2又は3に記載の方法であって、上記関心のあるターゲット画像パターンを含む複数のトレーニング画像を使って上記予備選択された基準データを計算し、それによって畳み込み核が形成されるステップを含む方法。

【0072】5. 上記4記載の方法において、上記予備選択されたデータを計算するステップ(100, 102, 104, 106及び108)が、上記トレーニング画像から引き出された平均化された画像から不要な画像特性を除去するために該平均化された画像の特定周波数成分をマスクするステップ(106)を含む方法。

【0073】6. 上記1, 2, 3, 4又は5記載の方法において、上記候補領域を選択するステップ(48)が、上記潜在的候補領域のそれぞれに関わる相関値を或1つの相関閾値と比較するステップ(118)を含む方法。

【0074】7. 入力画像(22)の中にターゲット画像パターンを検出する画像処理システム(10)であって、上記関心のあるターゲット画像パターンを含む確度を持つ上記入力画像中の画像領域を識別するため、上記入力画像と基準画像パターンとの相関をとり(114)、上記画像領域を選択する第1の選択手段(56)と、上記第1の選択手段と組んで動作し、上記ターゲット画像パターンを高い確度で含む候補領域(26, 32及び36)を選択するために上記画像領域を選択する第2の選択手段(66)で、候補領域を選択するために上記画像領域のグレースケール特性を調べる第2の選択装と、上記第2の選択手段に接続され上記候補領域を受け取り、上記候補領域を分析して、そ

れが上記ターゲット画像パターンを含んでいるかどうか検証する画像パターン検出部(14)と、を備える画像処理システム。

【0075】8. 上記7記載のシステムにおいて、上記第1の選択手段(56)は、上記入力画像に対して畳み込み核によって線形畳み込み操作を行うように構成された線形整合フィルタ(56)を含み、上記畳み込み核は上記基準画像パターンのひとつの数値表現であり、上記基準画像パターンは人間の顔を表すシステム。

【0076】9. 上記7又は8に記載のシステムであって、更に、複数のトレーニング画像から上記畳み込み核を生成するために上記第1の選択手段(56)と連携して動作する計算手段(16)を含み、この計算手段は、不要な画像特性に対応するトレーニング画像の特定の周波数成分を除去するためのマスカー(78)を含むシステム。

【0077】10. 上記7、8又は9記載のシステムにおいて、上記第2の選択手段(66)は、上記第1の選択手段(56)と直列に接続される非線形フィルタ(66)を含み、この非線形フィルタは、識別された各画像領域の部分領域から計算して得られたコントラスト値を成るコントラスト閾値と比較(122)することによって上記候補領域(26、32及び36)を選択するシステム。

【0078】

【発明の効果】この発明によると、速度と正確さにおいて顔などのターゲット画像の検出機能が向上する。

【図面の簡単な説明】

【図1】本発明による顔検出システムの構成を示すブロック図である。

【図2】図1の顔検出システムの流れに沿った色々な点での入力画像を描いたものである。

【図3】図1に示された候補選択部の内部構成を示すブ

ック図である。

【図4】図3の候補選択部によって調べられる8×8画像領域の部分領域を描いたものである。

【図5】図1に示された顔検出部の内部構成を示すブロック図である。

【図6】図1に示されたフィルタ核生成部の内部構成を示すブロック図。

【図7】図6のフィルタ核生成部によって行われるマスカー処理を描いた図。

【図8】本発明に従ってフィルタ核を計算する処理工程を示す流れ図。

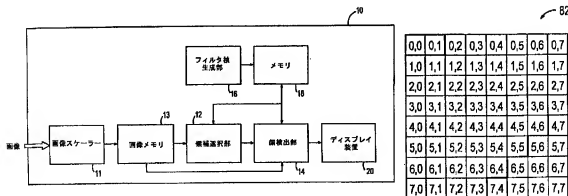
【図9】顔検出に際して、入力画像の中で潜在的に顔を含む画像領域を選ぶことによる入力画像の予備選択を本発明に従って行う処理工程を示すブロック図である。

【符号の説明】

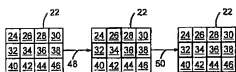
- 10 顔検出システム
- 12 候補選択部
- 14 顔検出部
- 16 フィルタ核生成部
- 52 線形相関部
- 54 処理モジュール
- 56 線形整合フィルタ
- 58, 60, 62 コントラスト計算部
- 64 決定モジュール
- 66 非線形フィルタ
- 72 ニューラル・ネットワーク
- 74 平均化ユニット
- 76 DCT部(離散コサイン変換部)
- 78 マスカー
- 80 IDCT部(逆離散コサイン変換部)

【図1】

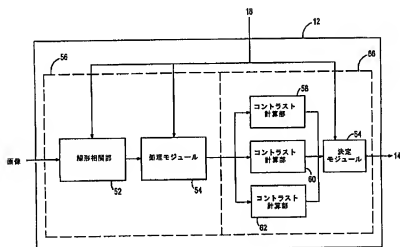
【図7】



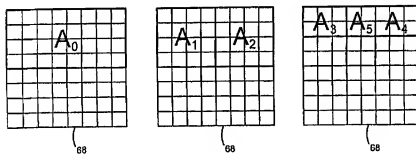
【図2】



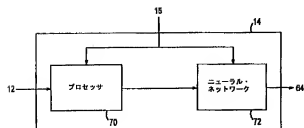
【図3】



【図4】

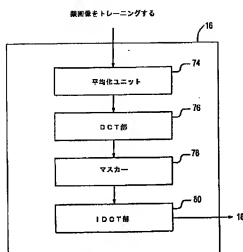


【図5】

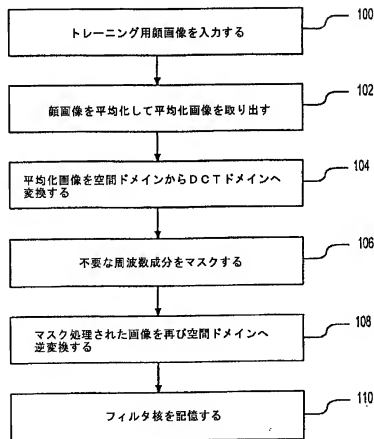




【図6】



【図8】



【図9】

